BANK OF CANADA
BANQUE DU CANADA

Staff Analytical Note/Note analytique du personnel 2018-19

The Scale and Scope of Online Retail

by Alex W. Chernoff

International Economic Analysis Department
Bank of Canada
Ottawa, Ontario, Canada K1A 0G9
achernoff@bankofcanada.ca

Bank of Canada staff analytical notes are short articles that focus on topical issues relevant to the current economic and financial context, produced independently from the Bank's Governing Council. This work may support or challenge prevailing policy orthodoxy. Therefore, the views expressed in this note are solely those of the authors and may differ from official Bank of Canada views. No responsibility for them should be attributed to the Bank.

Acknowledgements

The views in this paper are those of the author and do not necessarily reflect those of the Bank of Canada. I acknowledge financial support provided by the Network to Study Productivity in Canada from a Firm-Level Perspective, and Innovation, Science and Economic Development Canada. The firm-level analysis in this paper has been vetted by Statistics Canada and passed the compliance requirements of the agency's confidentiality protection and disclosure policy.

ISSN 2369-9639

© 2018 Bank of Canada

Abstract

This paper studies the growth of online retail over the period 1999–2012, using confidential firm-product-level data for Canada. The revenue of online retailers is decomposed into the contributions of product scope (the number of product categories) and product scale (average revenue per product category). During this period of rapid online sales growth, product scope dropped dramatically from an average of 59 product categories per firm in 1999 to 5 product categories per firm in 2012. Using data on Amazon.ca's expansion, I find suggestive evidence that this reduction in product scope may have been driven by increased competition in the online retail sector.

Bank topics: Firm dynamics; Digital currencies; International topics
JEL Codes: D22, L11, L81

Résumé

L'étude examine l'évolution du commerce de détail en ligne sur la période de 1999 à 2012 à l'aide de données confidentielles d'entreprises canadiennes ventilées par produits. Le chiffre d'affaires des détaillants en ligne est décomposé en deux éléments : la gamme, soit le nombre de catégories de produits, et l'ampleur, soit le chiffre d'affaires moyen par catégorie de produits. Sur cette période de croissance rapide des ventes en ligne, la gamme est passée, en moyenne, de 59 à 5 catégories de produits par entreprise. L'analyse de données sur l'expansion d'Amazon.ca donne à penser qu'une concurrence accrue dans le secteur du commerce de détail en ligne pourrait expliquer cette contraction de la gamme de produits.

Sujets : Dynamique des entreprises; Monnaies numériques; Questions internationales
Codes JEL : D22, L11, L81

1 Introduction

The long tail theory suggests that success in online markets is intrinsically linked to product diversity (Anderson 2004). In contrast, this paper shows that many firms have relied on product specialization to succeed in online markets. Using confidential firm-product-level data from Canada, I study the distribution of online retailers' sales across 115 product categories. I find that firms had sales in 59 product categories on average in 1999, and in only 5 in 2012. Using data on Amazon.ca's expansion into Canada, I find suggestive evidence that the reduction in product scope of Canadian online retailers may have been driven by increased competition.

2 Data and scale-scope decomposition

The firm-product-level data used in this paper comes from Statistics Canada's Annual Non-Store Retail Survey (RTNS), 1999–2012. This survey is designed to collect information on all Canadian retailers that do not operate brick and mortar retail stores. For my empirical analysis, I define a sample of "online firms" as the subset of RTNS firms that are classified under NAICS category 4541 (electronic shopping and mail-order houses), and have positive online sales.[1] For 2012, the aggregate total operating revenue of firms in my sample represents 51.0 per cent of the e-commerce sales, and 0.8 per cent of total sales, for the Canadian retail sector.[2]

The RTNS data include a commodity annex that details the distribution of each firm's sales across 115 different product categories. Drawing on the empirical literature on multi-product firms,[3]

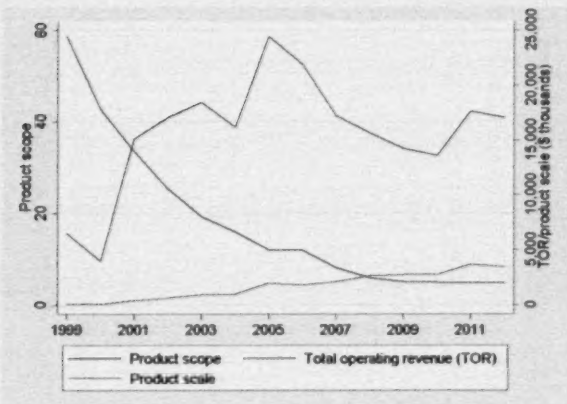
[1] For the purposes of the analysis, sales from electronic auctions are also included as online sales.
[2] Aggregate statistics for the Canadian retail sector (NAICS 44–45) were sourced from CANSIM table 080–0032.
[3] For example, an analogous scale-scope decomposition is used in Arkolakis and Muendler (2010).

I construct annual measures of product scope and scale as follows:

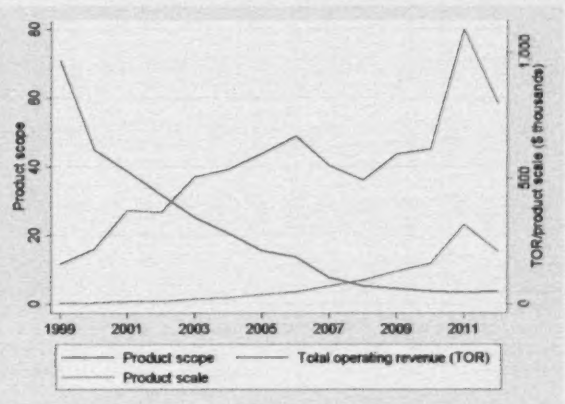
$$\overline{pq}_t = N_t \frac{pq_t}{N_t} \iff \underbrace{ln\left(\overline{pq}_t\right)}_{Revenue} = \underbrace{ln\left(N_t\right)}_{Product\ Scope} + \underbrace{ln\left(pq_t/N_t\right)}_{Product\ Scale}, \tag{1}$$

where $\overline{pq}_t$ is the average revenue of online retailers in year t; and N_t is product scope, defined as average count of product categories with positive sales. Product scale, pq_t/N_t, and product scope, N_t, are calculated across all firms, in each year from 1999–2012.

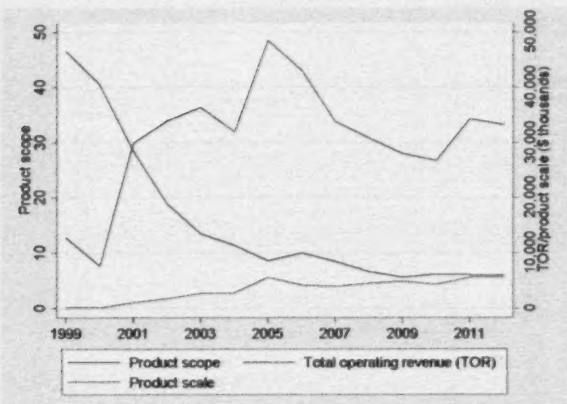
Chart 1: Decomposition of online retailers' scale and scope



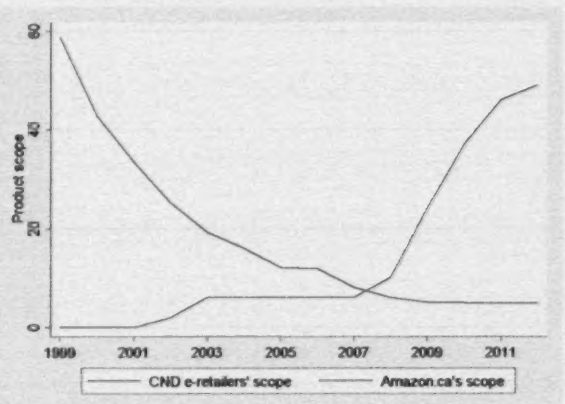
a. Scale and scope, full sample



b. Scale and scope, small firms: $TOR <= Median\ TOR$



c. Scale and scope, large firms: $TOR > Median\ TOR$



d. Product scope: Canadian online retailers and Amazon.ca

Chart 1 (panel a) plots the scale and scope decomposition in equation (1).[4] Product scope declined monotonically over the observation period, from 59 in 1999 to 5 in 2012. The average revenue of online retailers more than doubled over the observation period, growing from 6.5 million in 1999 to 17 million in 2012. With sales increasing and scope declining, product scale increased dramatically from 110,000 per product category in 1999 to 3.5 million per product category in 2012.

One potential explanation for the decline in product scope is that there may have been an increase in the number of small, single-product firms that were surveyed over time. To investigate this possibility, I recalculate the scale-scope decomposition in each year separately for the sub-samples of large and small firms. For the purposes of this robustness exercise, large firms are defined as the subset of firms with revenue above the annual median value, and small firms are those with revenue equal to or less than the annual median value. The results, presented in Chart 1 (panels b and c), show that the decline in product scope is apparent for small and large firms alike.

Research in the international trade literature suggests that increased competition encourages multi-product firms to specialize in their products of core competency (Mayer, Melitz and Ottaviano 2014). In Section 3, I use data on Amazon.ca's expansion into Canada to test the theory that increased competition contributed to the decline in product scope over the observation period. Using archived press releases from Amazon.com's *Press Room* and archived screen captures from Internet Archive Wayback Machine, I document the year of Amazon.ca's entry into the different product categories listed in the commodity annex of the RTNS. Chart 1 (panel d) illustrates that Amazon.ca's expansion in product scope sharply contrasts with the narrowing scope of Canadian online retailers over the period 1999–2012.

[4] A table with summary statistics for the variables used in the scale-scope decomposition is provided in the appendix.

3 Econometric model and results

The empirical strategy of this paper leverages the staggered expansion of Amazon.ca that is documented in Section 2. I use the following difference-in-differences estimating equation:

$$share_{j,t} = \delta_j + \alpha_t + \beta Amazon_{j,t} + u_{j,t}. \tag{2}$$

The dependent variable, $share_{j,t}$, is the share of Canadian online firms with positive sales in RTNS product category j in year t.[5] The key right-hand-side variable of interest is $Amazon_{j,t}$, which is a dummy variable that is 1 if Amazon.ca has entered product category j in year t, and 0 otherwise. The parameters δ_j and α_t are coefficients on product category and time dummy variables, respectively. To check for anticipation and persistence in the relationship between $Amazon_{j,t}$ and the dependent variable, I include an alternative specification that adds three lags and leads of the $Amazon_{j,t}$ variable to regression equation (2). In both specifications, the regressions are estimated by ordinary least squares with standard errors clustered at the product level.

The results are presented in Table 1. The first column presents the results from estimating the regression given by equation (2). The coefficient on the $Amazon_{j,t}$ variable is negative and statistically significant at the 1 per cent level. This result provides suggestive evidence that Amazon.ca's expansion, and increased online competition more generally, may partially explain the reduction in product scope of Canadian online retailers over the period 1999–2012. However, it is also likely that Amazon.ca coordinated its product expansion to target markets that were not already saturated by Canadian firms. As such, I acknowledge that reverse causality may partially explain the negative

[5]The variable $share_{j,t}$ is calculated using a custom tabulation prepared by Statistics Canada at the request of the author. This custom tabulation includes the count of firms in each RTNS product category in every year over the period 1999–2012. Counts were not released for product categories with two or fewer firms to meet the compliance requirements of the agency's confidentiality protection and disclosure policy. In a limited number of cases, product categories with three firms were also not released in accordance with the agency's confidentiality protection and disclosure policy. Product categories where the counts were not released are treated as missing in estimating equation (2).

and highly significant coefficient on the $Amazon_{j,t}$ variable. For this reason, the results should be interpreted as estimates of conditional correlations rather than causal effects.

Table 1: Difference-in-differences regression results

	1	2
$Amazon_{j,t+3}$		-0.0657***
		(0.0115)
$Amazon_{j,t+2}$		-0.00667
		(0.00592)
$Amazon_{j,t+1}$		-0.00910**
		(0.00369)
$Amazon_{j,t}$	-0.0564***	-0.0150***
	(0.0113)	(0.00503)
$Amazon_{j,t-1}$		-0.00709
		(0.00454)
$Amazon_{j,t-2}$		-0.0174***
		(0.00439)
$Amazon_{j,t-3}$		-0.0217*
		(0.0128)
Number of lags	0	3
Number of leads	0	3
Product and year dummy variables	Yes	Yes
Observations	1,485	1,485
R-squared	0.915	0.921

Note: Standard errors (in parentheses) are clustered at the product level; *** p<0.01, ** p<0.05, * p<0.1.

The results in the second column of Table 1 show evidence of statistically significant anticipatory and persistent effects from Amazon.ca's expansion. In particular, the coefficient on $Amazon_{j,t+3}$ is negative, statistically significant, and larger in magnitude than any of the other lags and leads. One interpretation of this result would be to infer that Canadian firms anticipated Amazon.ca's product expansion, and aggressively exited these product categories three years in advance. However, it is unlikely that Canadian firms had such advance knowledge of Amazon.ca's expansion plans. A more probable explanation is that Canadian online firms may have been responding to Amazon.com's expansion, or other changes in the online market environment that predated Amazon.ca's

expansion.[6] The lagged Amazon variables show evidence of a non-linear relationship between Amazon.ca's product expansion and the dependent variable. Relative to the magnitude of the coefficient on the $Amazon_{j,t}$ variable, the coefficient on the first lag is smaller and not statistically significant, while the second and third lags are larger and statistically significant at the 1 and 10 per cent levels respectively.

4 Conclusion

The results of this paper show that the average product scope of Canadian online retailers declined dramatically over the period 1999–2012. At first glance, this result seems to contradict the findings of the long tail literature, which has shown that product diversity is a key to success in online retail. However, an important distinction between the long tail literature and my analysis is the level of product aggregation. This paper defines product scope at a relatively coarse level of product aggregation (i.e., the RTNS has 115 product categories). By comparison, empirical studies on the long tail typically use highly disaggregated product data, such as stock keeping unit (SKU) data (Brynjolfsson, Hu and Smith 2010). This difference in aggregation is important, as there are many examples of online firms that have narrow product scope yet offer a wide range of varieties to customers. For example, Netflix and Spotify specialize in digital video and music respectively, and offer consumers an abundance of choices within these broadly defined product categories. To the best of my knowledge, this paper is the first to highlight the increasing prevalence of product category specialization among online retailers.

[6] Many Canadian consumers purchase directly from Amazon.com. Therefore, competition in the Canadian online retail sector may have been affected by Amazon prior to the establishment of its Canadian subsidiary, Amazon.ca. Ebay.ca was established in 2000 and also likely affected the market environment in the Canadian online retail sector.

References

Anderson, Chris (2004) 'The long tail.' *Wired Magazine* 12(10), 170–177

Arkolakis, Costas, and Marc-Andreas Muendler (2010) 'The extensive margin of exporting products: a firm-level analysis.' *NBER Working Paper* 16641, 1–50

Brynjolfsson, Erik, Yu (Jeffrey) Hu, and Michael D. Smith (2010) 'Research commentary - long tails vs. superstars: the effect of information technology on product variety and sales concentration patterns.' *Information Systems Research* 21(4), 736–747

Mayer, Thierry, Marc J. Melitz, and Gianmarco I. P. Ottaviano (2014) 'Market size, competition, and the product mix of exporters.' *American Economic Review* 104(2), 495–536

Appendix

Table A-1: Summary statistics for variables used in scale-scope decomposition in Chart 1 (panel a)

Variable	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012
pq_t	6,513	3,986	15,106	17,035	18,409	16,145	24,415	21,842	17,172	15,633	14,219	13,580	17,596	16,992
	(2,209)	(1,091)	(8,191)	(8,925)	(6,561)	(5,503)	(8,781)	(8,552)	(6,459)	(5,461)	(4,419)	(3,791)	(4,214)	(3,668)
N_t	59	43	34	25	19	16	12	12	8	6	5	5	5	5
	(3.63)	(2.24)	(2.74)	(2.61)	(2.13)	(1.86)	(1.77)	(1.60)	(1.10)	(0.79)	(0.66)	(0.58)	(0.54)	(0.54)
Observations	144	252	186	175	190	219	167	180	208	218	223	243	219	248

Note: Mean values are reported with standard errors beneath in parentheses. Average revenue, pq_t, is in thousands of dollars. Product scope, N_t is the average count of product categories with positive sales, rounded to the nearest integer.